REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
17-03-2005	Final Report	JULY 2004 to JULY 2005

4. TITLE AND SUBTITLE
Exploring the Readability of Consent Forms in Human Research in the United States Army

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Captain Heidi P. Mon

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)

U.S. Army-Baylor Program
3151 Scott Road
Fort Sam Houston, Texas 78234

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Medical Department Center and School
BLDG 2841 MCCS-HRA (Army-Baylor Program in Healthcare Administration)
3151 Scott Road, Suite 1411
Fort Sam Houston, Texas 78234-6135

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
35-05

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; Distribution is unlimited

13. SUPPLEMENTARY NOTES
Prepared with the cooperation of Captain Clemons S. Kruse

14. ABSTRACT

Informed consent documents used in human subject research within the United States Army appear increasingly complex and lengthy and are rife with medical and legal terminology. The intent of these consent forms becomes blurred between whether their primary purpose is to inform the patient or to protect the researcher and organization from litigation. A literature review highlighted two observations: (a) consistently, every article published about consent forms concluded that these documents were too complex for the layperson; and (b) there is a gap in the literature concerning the readability of consent forms in military protocols. Using a 1997 study conducted by Mader and Playe (n = 94) as a foundation, this study evaluated the readability of consent forms (n = 60) in human research performed within the United States Army. Studying the effects of ten dependent variables based on two levels of risk (minimal risk and greater than minimal risk), five of the ten variables were found significant ($p<.01$). The results demonstrate that the readability of consent forms within the U.S. Army is too complex for the average reader. A readability standard of the sixth grade level would better serve the interests of participants in human subject research.

15. SUBJECT TERMS
Readability, CFR, Ethics, Consent Forms, Autonomy

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Education Technician
Unclassified	Unclassified	Unclassified	SAR	56	19b. TELEPHONE NUMBER *(include area code)* (210) 221-6443

Running head: READABILITY OF CONSENT FORMS IN HUMAN RESEARCH

Exploring the Readability of Consent Forms in Human Research in the United States Army

Great Plains Regional Medical Command, Fort Sam Houston, Texas

CPT Heidi P. Mon

U.S. Army-Baylor University Graduate Program in Health Care Administration

March 2005

ACKNOWLEDGEMENTS

I would like to extend my heartfelt gratitude and appreciation to those who have directly influenced and participated in the completion of this study. Particularly, I would like to thank my husband, MAJ Robert Mon, for his enduring love and support, COL Lee Briggs for his support and guidance, COL Scott Martin and COL (R) James Lamiell for their clinical expertise and for providing the documents utilized in this study, Dr. Karin Zucker for her literary acumen and guidance, CPT Scott Kruse for conducting this study with me, and my family and fellow Baylor classmates for their support.

Abstract

Informed consent documents used in human subject research within the United States Army appear increasingly complex and lengthy and are rife with medical and legal terminology. The intent of these consent forms becomes blurred between whether their primary purpose is to inform the patient or to protect the researcher and organization from litigation. A literature review highlighted two observations: (a) consistently, every article published about consent forms concluded that these documents were too complex for the layperson; and (b) there is a gap in the literature concerning the readability of consent forms in military protocols. Using a 1997 study conducted by Mader and Playe (n = 94) as a foundation, this study evaluated the readability of consent forms (n = 60) in human research performed within the United States Army. Studying the effects of ten dependent variables based on two levels of risk (minimal risk and greater than minimal risk), five of the ten variables were found significant ($p<.01$). The results demonstrate that the readability of consent forms within the U.S. Army is too complex for the average reader. A readability standard of the sixth grade level would better serve the interests of participants in human subject research.

Table of Contents

List of Tables

Table

List of Figures

Figure

Introduction

The introduction, literature, purpose and procedure sections of this study have been co-authored with CPT Clemens S. Kruse. These sections are mirrored exactly as written here in his study.

During our second semester of the U.S. Army Baylor Program, CPT Kruse and I conducted an independent study under the direction of Dr. Karin Zucker, Associate Professor. This study involved performing a comprehensive literature review of the readability of informed consent documents in the conduct of human research within the Department of Defense (DoD). While an abundance of literature exists on readability of informed consent documents in the civilian sector, there was little found involving DoD. The lack of literature prompted this current study.

In order to compare the results involving the United States Army and Air Force informed consent documents and to verify the validity and reliability of the comparison, the study conducted on the respective service's consent forms had to be identical. Therefore, CPT Kruse and I co-authored the purpose and procedure sections which now appear in our respective studies. Further, we maintained communications throughout our respective studies to ensure that any difficulties encountered during the performance of the studies were handled in the same fashion.

Conditions That Prompted the Study

When we attended the Institutional Review Board (IRB) at Brooke Army Medical Center (BAMC) we found the consent forms included with the protocols were complex and lengthy. When the risk to a subject was high, the complexity and length of the consent forms seemed to increase. All were rife with both medical and legal terminology. It became difficult to discern

whether the intent of consent forms was to inform the patient or protect the researcher and organization. A brief literature review highlighted two observations: (a) consistently, every article published about consent forms concluded that these documents were too complex for the layperson; and (b) there is a gap in the literature concerning the readability of consent forms in military protocols. A comprehensive literature review ensued on the readability of consent forms in military, human research studies.

Statement of the Problem or Question

The primary question was, "What is the readability of consent forms in military, human research studies?" To answer this question, it was necessary to: (a) operationally define readability, (b) discern the intent of the consent form in human research, and (c) explore the ethics inherent to this subject. Overall, the observation was that the reading level of consent forms is too high. Reading levels above the average person's ability do not facilitate his/her understanding of the research procedures and their risks, benefits, and alternatives to the research procedures. Supporting studies to this claim include Glazer-Waldman, Hall and Weiner's research (1985) that demonstrated 40% of adults tested at a Texas hospital read below the 6th grade, and Ott and Hardie's study (1997) that suggested written materials given to patients should not be above the sixth grade level. Using the previously mentioned studies and the guidance provided in Army Regulation 40-38 (The Clinical Investigation Program) that consent forms "will be written in language that is easily understandable by the subject," the average person's ability is defined as sixth grade for purposes of this study (Army Regulation 40-38, 1989, p. 4).

Literature Review

The History of Informed Consent

Scrutiny of human subject research exposes a sinister side to medical research and a long history of grossly unethical experiments performed on non-consenting patients, even though its regulation reaches back to World War II when the horrific Nazi experimentation was exposed. Following the Second World War, the United States tried and executed a number of involved Germans for war crimes and crimes against humanity in what became known as The Doctors' Trial at Nuremberg. (*United States v. Karl Brandt*, 1947). The opinion in that case included 10 basic principles for human research, called the Nuremberg Code. Thereafter, Article seven of *The Universal Declaration of Human Rights* was passed to protect research subjects from torture, and cruel, inhumane treatment (The United Nations General Assembly, 1948). Later, the World Medical Association published *The Declaration of Helsinki (1964)* that safeguarded the health of the subjects (as cited by Zucker & Boyle, 2000). Finally, the *Belmont Report* (1976) stood as ethical principles and guidelines for the protection of human subject research. Together, the Nuremburg Code of 1947 and the Declaration of Helsinki of 1964 formed the basis of United States federal regulations that govern federally supported research with human subjects (Woodward, 1999). Both codes demanded that the rights of individual patients and human research subjects be placed above scientific and societal goals. Yet, the experimentation without informed consent continued.

Several notorious cases of unethical human experimentation tarnish America's rich history of medical advancements. The Tuskegee Experiment, from 1932 through 1972, involved 399 unknowing African American participants in a study involving the effects of untreated syphilis (Jones, 1993). In 1952, Harold Blauer was subjected to injections of mescaline derivatives

supplied by the U. S. Army Chemical Corps. The purpose of the injections was to determine the effects of chemical agents on humans, but they were administered to Mr. Blauer under the guise they would cure his depression (Albarelli & Kelly, 2001). In 1953, without parental consent, a premature infant was given a high dose of oxygen as part of an experiment. The infant went blind (Standler, 1997). In 1963, 22 chronically ill non-cancer patients unknowingly received intradermal injections of live human cancer cells. The experiment's purpose was to learn if foreign cancer cells would survive longer in incapacitated non-cancer patients than in patients debilitated by cancer (Standler, 1997). In 1964, personnel at the Willowbrook State Hospital in New York injected severely retarded children with hepatitis virus. The parents 'consented' to the injections believing they were vaccinations (University of Utah, 2004c). From 1960 to 1972, cancer patients in Cincinnati were exposed to large doses of whole body radiation as part of an experiment, although they thought they were receiving standard treatments. Several died prematurely as a result of radiation exposure (University of Utah, 2004a).

In 1974, the National Research Act established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. During the next 4 years, the commission identified the basic ethical principals that should underlie the conduct of biomedical and behavioral research involving human subjects. Additionally, it recommended guidelines to ensure that the research was conducted in accordance with those principles. On September 30, 1978, the commission submitted a report defining the basic ethical principles in human subject research titled The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research (Beauchamp & Childress, 2001).

The Belmont Report set forth the three requirements essential for the ethical conduct of research involving human subjects: autonomy (respect for persons), beneficence, and justice.

The report also defined how these principles apply to the conduct of research. The principle of autonomy (respect for persons) underlies the need to obtain informed consent (Beauchamp & Childress, 2001). Informed consent provides a primary means by which federal regulations pertaining to human subject research seek to protect the rights and welfare of research subjects (Woodward, 1999).

Informed consent includes three elements: information, voluntariness, and comprehension, (U.S. Department of Health and Human Services, 2004). Enough information must be provided to potential research subjects for them to decide whether to participate in the research. Elements of essential information from Volume 45 of the Code of Federal Regulation, Part 46, Public Welfare, (known as *The Common Rule)* include: the purpose of the research purpose, expected duration of the subject's participation, description of procedures and which procedures are experimental, description of reasonably foreseeable risks or discomforts to the subject, potential benefits, alternative procedures that might benefit the subject, extent of confidentiality, explanation of compensation, a point of contact for additional questions, a point of contact if injury occurs, a statement of voluntariness, a statement of reassurance that failure to participate will not cause penalty or loss of benefits to which the subject is otherwise entitled, and a statement of reassurance that the subject can discontinue participation at any time without penalty or loss of benefits (Public Welfare, 1991). Consent to participate in research must be completely voluntary in nature and free from coercion. Finally, study participants must be able to comprehend the information presented to them. "The presentation of information must be adapted to the subject's capacity to understand it; testing to ensure that subjects have understood may be warranted" (U.S. Department of Health and Human Services, 2004).

An abundance of literature urges researchers to write simple and brief consent forms, yet

consent forms range in complexity from grade 6 through grade11 and beyond. Although the process of informed consent involves more than the written consent form, the basics of the research are first defined in the consent form; and, thus, it forms the basis for the potential participant's ability to comprehend the purpose, procedure, risks, benefits and alternatives risks, and then volunteer for participation. The readability of a consent form is vital to obtaining informed consent in human subject research.

Assessing Reading Level

The term readability refers to all the factors that affect success in reading and understanding text including the interest level and motivation of the reader, the legibility of the print, and the complexity of words and sentences in relation to the reading ability of the reader (Johnson, 2004). The determination of readability addresses the problem of matching individual reading levels to the difficulty of the text.

Several tests exist to assess readability or reading level. The primary purpose of these tests is to provide an assessment of the density of the text. The Gunning Fog Index uses the number of words per paragraph, the number of sentences per paragraph, and the number of words with three syllables or more to determine the number of years of education needed by the reader to understand the text. Shorter sentences written in plain English score better than longer sentences written in complicated language (Gnome, 2004).

The Flesch-Kincaid Formula assesses grade level and reading age by determining the average sentence length and the average number of syllables per word. Similar to the Flesch-Kincaid Formula, the McLaughlin Simplified Measure of Gobbledygook (SMOG) readability formula computes readability based on the average number of syllables per word and the average number of words per sentence. However, the SMOG formula computes a reading level for

written materials that is not associated with a grade of school such as that calculated by the Flesch-Kincaid Formula (University of Utah, 2004b). Additionally, the McLaughlin Formula tends to calculate higher values than other readability formulas because this test intends to predict the level necessary for 100% comprehension of the text (Johnson, 2004).

The Fry Readability Graph uses the average number of sentences and the average number of syllables per 100-word passage. These averages are then applied to the Fry graph to determine reading age in years. The Powers-Sumner-Kearl Formula is most suitable for analysis of material for 7 to 10 years old readers, and it uses the average sentence length (number of words /number of sentences) and the number of syllables per 100 words to determine reading age. The FORCAST Formula was specifically designed for assessing the readability of U.S. Army technical manuals. As such, it is not suitable for primary age reading material (Johnson, 2004). This formula does not require full sentences to access readability. Grade level is calculated by dividing the number of single-syllable words in a 150-word passage by 10. This number is then subtracted from 20. Reading age is determined similarly by subtracting the number of single-syllable words divided by 10 from 25 (Johnson, 2004).

There are several limitations to assessing reading level by any readability test, however. First, a readability test predicts the 'break-off' point for a reader of a specific reading age (Johnson, 2004). If a reading level is measured at 10^{th} grade, an average 10^{th} grader would be at the upper limit of his/her reading comprehension. Most readability formulas are based on a 50% correct answer score in a comprehension test (the McLaughlin SMOG formula is an exception). If a reading level of 10 years was predicted, an average 10-year-old student would only score 50% on a test of comprehension of that text (Johnson, 2004). Readability tests alone may not be the only evaluator of the suitability of text, which is another limitation. Other factors may need

to be considered such as the size of type and length of line, sentence structure, the number of words per page, the use of color, the use of diagrams, the page layout, and the use of space between paragraphs (Johnson, 2004).

The concept of readability is based on "functional literacy" (Lee, 1999). Individuals not only need to be able to read, but also to understand and act on that understanding, especially when considering the risks and benefits of participation in a human subject research study. In response to the scrutiny of readability of patient material, three specific tests were developed within the last few years to evaluate medical literacy. The Rapid Estimate of Adult Literacy in Medicine (REALM) was designed for use in public health and primary care settings to identify patients with low reading levels. A second test is the Test of Functional Health Literacy in Adults (TFHLA). This test more fully assesses functional literacy as well as reading ability. Analyses indicate that results of this test correlate with scores on more generalized reading tests.

Readability of Informed Consent Documents

The Declaration of Helsinki requires human researchers to "adequately inform" participants concerning the trial's aims, methods, expected benefits, risks, and alternatives. Unfortunately, the authors of the Declaration failed to define the elements of adequate information. The writers also did not describe the end state of being adequately informed. Informed consent received considerable attention by Beauchamp and Childress (2001). These icons of medical ethics defined informed consent as "an autonomous authorization of individuals of a medical intervention or of involvement in research" (p. 78). Meisel and Roth (1981) and the Belmont Report (1976) posit (as referenced by Beauchamp & Childress, 2001, p.79) two of the elements of informed consent are *information* and *consent*. Information is not merely disclosure of information but is also comprehension of what is disclosed (Beauchamp & Childress, 2001).

Consent is more complicated. This latter element consists of five elements: (a) competence, (b) disclosure, (c) understanding, (d) voluntariness, and (e) consent (p. 79). These building blocks create a pyramid of consent, the absence of which makes the structure unstable. "One gives and informed consent to an intervention if (and perhaps only if) one is competent to act, receives a thorough disclosure, comprehends the disclosure, acts voluntarily, and consents to the intervention" (p. 79).

The rules and regulations of the U.S. Department of Health and Human Services instruct the authors of consent forms to write these documents using language that is understandable by the subject; the Public Welfare Title of the Code of Federal Regulation (CFR), however, does not specify a readability standard by established indices (Public Welfare, 2004). In order for a consent form to adequately inform a participant, the consent form must use language commensurate with his ability to read and comprehend. Pursuant to this intent, researchers such as Mader and Playe (1997) explored standards set by institutional review boards (IRBs). Esty, Musseau, & Keehn (as cited in Mader and Playe, 1997) claim a preponderance of IRBs interpret the Code of Federal Regulation's instruction as a readability scale no higher than the sixth grade.

Ferguson (2002) explored medical trial participants' perceptions of the adequacy of the information they were provided and their understanding of this information. Participants in Ferguson's study felt they understood the experiment's intent, methods, benefits, risks, and alternatives. The participants even felt they had adequate time to ask questions. When they were questioned about the study, however, the depth of their understanding was shallow. Ferguson referenced Howard and DeMets' (1981) findings that, "research subjects . . . do not adequately understand the programs involved" (p. 48). Researchers focus on providing information, but few seek to ensure that the participants understand what they were provided (Ferguson, 2002).

Arthur (1995) explored the effects of repeated exposure to medical information by providing an additional pamphlet to patients upon their discharge to increase the frequency of their exposure to the details of the experiment. She found a statistically significant increase in recall of the medical information concerning specific conditions and medications. The research of Ferguson and Arthur demonstrated that repeated exposure to the information in consent forms might increase the participants' recall of the information, but not necessarily their level of understanding. The first principle in the Nuremburg Code (as referenced by Zucker, 2000, p. 845) requires that the participant "should have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision." Because participants must be able to read and understand the details of the experiment, a readability assessment of the consent form should be used as a measurement tool.

The seven readability assessments in the literature are not immune to criticism. Some researchers question the validity of the readability tests. Others challenge the readability thresholds set by tests. Furthermore, different researchers set dissimilar readability thresholds, and their range is wide. The science behind the selection of readability parameters is not exacting in nature.

Three studies exemplify the wide range of readability thresholds. Ott and Hardie (1997) cited the Flesch reading ease score as the U.S. Government standard for military documents and specified its readability goal of seventh to eighth grade (based on a readability score between 60 and 70). Despite their reference to the U.S. Government standard, Ott and Hardie set their readability threshold at the sixth grade. Arthur (1995) evaluated the readability of medical pamphlets in the United Kingdom. His literature review expressed the importance of

discernment and caution when interpreting the results from various readability assessments. He found that the algorithms used by the tests are complex and can render a wide range of reading ages. Using the Flesch, FOG, and SMOG indices, Arthur set an acceptable readability level at 11.6 years of schooling. Mader and Playe (1997) set their readability goal at fifth grade but offered little justification for their choice.

How low a researcher should set his readability-criteria threshold in order to ensure a high percentage of adequately informed participants is still unclear. Researchers should select thresholds commensurate with their audience. Arthur (1995) found many pamphlets written at a readability score of 15. Clearly, this threshold is too high for most readers. Glazer-Waldman, Hall, and Weiner (1985) found 40% of adults at a Texas hospital read below the sixth grade level. Ott and Hardie cited research by Walmsley and Allington (1982) that found 33% of elderly adults at a New York senior center read below the fourth grade level and 35% read between the fifth and eighth grade level. Although Mader and Playe (1997) set their readability goal at fifth grade, they found the readability average of the medical material they evaluated was above a 10^{th} grade level. Almost universally, Walmsley and Allington (1982), Mader and Playe (1997), Glazer-Waldman, Hall, and Weiner (1995), Ott and Hardy (1997), and Ferguson (2002) agreed that the consent forms they evaluated were written at a level above the participants' ability to comprehend their message. Such conclusions question the ability of most consent forms to adequately inform participants of risks, benefits, and alternatives.

Another common criticism of the various readability assessments is that researchers cannot equitably compare their results without a baseline understanding of the indices. Mader and Playe (1997) assessed readability using Right-Writer 5.0, which is a program that checks documents for grammar and spelling. This program is comprised of three indices: Flesch-

Kincaid, Flesch, and Gunning Fog. These indexes all provide readability levels that have become industry standards, but each bases its conclusion on distinct algorithms. The resulting readability level for each test cannot necessarily be compared with the results of the other two. Other researchers use different readability programs that calculate similar indices. Ott and Hardie (1997) evaluated the readability of advance directives using another program, similar to Right-Writer 5.0, Grammatik II. This program also calculated the readability scores of Flesch, Flesch-Kincaid, and Gunning Fog indices. Their study evaluated the scores against each other. According to the Grammatik II program results, the Flesch and Gunning Fog indices consistently provide higher readability scores than the Flesch-Kincaid index. Ott and Hardie did not interpret the results, provide reasons for the difference, or suggest one test over the others. Instead, the researchers left such conclusions to the reader. Such inconsistent results reinforce concerns about the reliability of the tests.

Researchers complain that they must provide a vast amount of information to a population that will most likely not be able to understand it (Ferguson, 2002). They must fulfill the requirements of the Nuremburg Code, the Universal Declaration of Human Rights, the Declaration of Helskini, and the Belmont Report. The complexity of the requirements creates the potential for an intricate and complex document.

Ferguson (2002) highlighted the bifurcated role that researchers must play by describing the direct relationship between the extensive nature of the consent process and the resulting satisfaction of the participants. Participants appreciated the extent of the information and felt it was necessary for their understanding (Ferguson, 2002). Whether the information increased their understanding of the consent material was unknown. Further studies are needed to assess the validity of participants' perception of their understanding (Ferguson, 2002). Ferguson's research

stresses the importance of conveying a complete message to the participants of human research during the consent process. A consent form authored with an appropriate level of readability enables the participant to better understand the elements and effects of a study. A participant that fully understands a study helps the researcher meet the Declaration of Helsinki's requirement of *adequately informed consent.*

Relying on a report from the Health Journal of Family Practice which stated that "almost half of American adults read at or below the 8th grade level" (1988), members from a group of IRBs developed a set of consent form templates for researchers to use (Paasche-Orlow, Taylor, & Brancati, 2003). These templates ranged in readability from 4th grade to college level and were developed to assist researchers in writing consent forms at a level that most participants can understand. The IRBs provided these templates to medical schools and research institutes (Paasche-Orlow, Taylor, & Brancati, 2003). They serve as an excellent resource today for researchers trying to simplify the language of their consent forms. A combination of these templates and common readability assessments should provide researchers a tool that will allow them to improve readability and comprehension. Improved readability should enable participants to better understand the details of the study, benefits, risks, and alternatives.

Autonomy and MHS Protocols

Beauchamp and Childress (2001) require five elements to satisfy the bioethical tenet of autonomy. They require liberty, which is the "independence from controlling influence," and agency, which is the mental "capacity for intentional action" (p. 58). The other three required elements, inherent to respect for autonomy, explain that *normal choosers are* those who act "(1) intentionally, (2) with understanding, and (3) without controlling influences that determine their action" (p. 59).

The vast amount of literature that discusses the need to improve the readability of consent forms begs the question: Why do researchers continue to author consent forms far above the readability level of average participants? Exploring that question is beyond the scope of this study. A more focused question for this study is: How widespread is the problem within the military health system (MHS)?

The Department of Defense (DOD) conducts a large amount of human subject research every year. 10 USC 980 requires that funds appropriated to the Department of Defense may not be used for research involving a human being as an experimental subject unless (1) the informed consent of the subject is obtained in advance; or (2) in the case of research intended to be beneficial to the subject, the informed consent of the subject or a legal representative of the subject is obtained in advance. The Secretary of Defense can waive these requirements with respect to a specific project if the project's purpose is to advance the development of a medical product necessary to the armed forces and if the research project may directly benefit the research subject and is carried out in accordance with all other applicable laws.

DOD human research studies solicit participants from the MHS community, to include retirees and trauma patients brought into MHS emergency rooms. Human subject research within the Department of Defense is divided into minimal risk studies and greater than minimal risk studies. Minimal risk studies, as defined in Part 219 of Volume 32 of the Code of Federal Regulations, National Defense, are studies where the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests (National Defense, 2004). Greater than minimal risk studies are those outside the studies defined as minimal risk. Degree of risk is established by following the

guidelines in 32 CFR Part 219 and Army Regulation 40-38. AR 40-38 (1989) instructs investigators to author volunteer agreements "in language that is easily understandable" (3-5.(6).c.3), language that is not otherwise defined.

Military consent forms written to a participant from a military community should adopt a readability standard, such as that used in Mader and Playe (1999). A readability standard of the sixth grade level is difficult to meet, but such a standard would better serve the interests of the participants. This study explores the readability of consent forms for human subject research studies of minimal and greater than minimal risk conducted by MHS researchers in active duty military treatment and research facilities.

Intent of Consent Forms

The intent behind a reasonably understandable consent form is to enable the participant to weigh the benefits against the risks and alternatives inherent in the research design. This decision process is necessary to empower the participant with autonomy (Luce, 2003). Because "every human being of adult years and sound mind has the right to determine what shall happen to his body" (*Schloendorff v. Society of New York Hospital*, 1914), the consent form plays an integral role in the consent process.

It is worthy of note that in most cases, the consent form is not the primary method of informing the participant of the details of the study and the inherent benefits, risks, and alternatives – it is certainly not the sole means. The consent form is combined with an interview, a question and answer period, and often a video – all, not infrequently, followed by additional face-to-face discussions during the consent process. Researchers should attempt to tailor the entire consent process, particularly the readability of the consent form, to their audience. The participants' ability to comprehend may be reduced by medical conditions. Participants in

psychiatry or oncology studies could be particularly vulnerable in this regard. Not only is each patient adjusting to a potentially life-altering sickness, but he must also endure a consent process laden with complex medical and legal terminology.

Oncology consent forms are inherently lengthy. The authors of oncology consent forms must satisfy a group of stakeholders, which includes the hospital attorney, the researcher himself, and the members of the IRB. The legal review balances due diligence with institutional protection. The researcher himself weighs anonymous advancement of his science with his desire for recognition and advancement in his field. The IRB weighs the risks of the trial with the potential benefits. The ethical intent of the research should be to use the consent form to facilitate autonomy, and this intent should rest equally on all shoulders.

Luce (2003) questioned the applicability of the consent process in deference to the psychological state of critically ill patients. He explored the legal competence of the critically ill and discussed the absence of legal surrogates. If a patient is otherwise competent, does his critically ill status alter his ability to make decisions on his own behalf? Does the mental state subsequent to a grim diagnosis and dim prognosis of life expectancy create in the patient indifference to risk in light of remote benefit?

Patients automatically assume the physician has their best interests at heart. Many doctors do have such altruistic motives, but the few who do not raise several questions. At what point do professional notoriety and advancement and the possibility of monetary gains change altruism into self-interest? At what point does a seasoned researcher become aware of his changing motives and sense ambivalence? Does such extreme self-interested motives affect the research, the consent process, or the participants' autonomy?

The literature suggests a consistent trend in the readability of consent forms. The psychological state of critically ill patients that creates indifference in their decision-making process may perpetuate this trend by reinforcing a sloppy consent process. If the patients' desperation supports the researchers' ambition, are the ethics of the situation compromised? Is the emphasis on autonomy as defined by Beauchamp and Childress not-applicable in palliative care? If there has been no improvement in the readability of consent forms in 30 – 40 years, is it because the medical community has not focused on the issue, or is it because many patients do not care about the risks involved in a study if there is even the possibility of only a modicum of benefit?

If a researcher authors a consent form above an acceptable level of readability, is he abiding by the Declaration of Helsinki's requirement to adequately inform participants? Perhaps he is, if the complicated consent form is adequately explained during the consent process.

The results of this study support the trend noted in the literature, i.e., that consent forms for human research are written above the level of comprehension of the average participant. Researchers' may compensate for the imbalance of readability with complexity in the rest of the consent process, or it may not be.

Matot, Pizov, and Sprung (1998) studied the legitimacy of the human research process. The Common Rule requires that anybody who receives money from the federal government to perform human subject research must follow the Department of Health and Human Services published regulations for the protection of human subjects (Public Welfare, 1991). Though the Common Rule requires the IRB process for human research (Zucker & Boyle, 2000), Matot, Pizov & Sprung found that 41% of the 279 research studies they reviewed involving critically ill patients were either not reviewed by an IRB or the issue of informed consent was not addressed

(1998). Though the Declaration of Helsinki compels medical journals to decline to publish research without IRB approval or informed consent, many journals still publish the research (Matot, Pizov & Sprung, 1998). If the Nuremberg Code, the Declaration of Helsinki, the Belmont Report, and the Common Rule require adequate informed consent, why has there been no improvement in level of readability of consent forms? Perhaps the reason for this trend transcends blind, generational mentoring. Could a justification for complex consent forms stem from a medical professional's desire to advance in his field? Professional associations such as the American College of Healthcare Executives, the American Association of Medical Assistants, and the Association of Medical Surgeons of the United States facilitate the dissemination of information and networking. A professional's ability to attain name recognition largely originates with publication in widely read, peer-reviewed journals. Would a more exhaustive, simpler-to-understand consent form enable a participant to properly weigh the dismal probabilities of benefit against the high probability of harm? If this Spartan message were conveyed to the participant, would the researcher find sufficient numbers for statistical significance? Does a person's desperation become the deciding factor for participation in a Phase I study? Does desperation replace reason when evaluating the study's risks and benefits? Will the researchers in Phase I studies ever see a decrease in participation? If research participants became more reticent about participating in medical research trials, would research institutes be able to process the volumes of data necessary for future funding? Perhaps self interest fuels the narcissistic motivation for a sloppy consent process.

Purpose (Variables/Working Hypothesis)

This study examines the readability statistics based on the risk or potential harm to a human research subject. The two risk categories in military human subject research, as defined

by the Common Rule and Army Regulation 40-38, are minimal risk and greater than minimal risk. The readability statistics are calculated using Microsoft Word™, which utilizes the Flesch-Kincaid criteria. Consent form readability variables includes the Flesch Reading Ease, Flesch-Kincaid grade level, number of words per document, number of characters per document, number of paragraphs per document, number of sentences per document, average number of sentences per paragraph, average number of words per sentence, average number of characters per word, and number of passive sentences per document. These statistics are quantitative in nature, enabling statistical analysis on the Statistical Package for the Social Sciences (SPSS™) version12.0. The alternate hypothesis is that the readability of consent forms in military protocols is sufficient to adequately inform the average military reader (Flesch-Kincaid grade level of 6). The null hypothesis is that the readability of consent forms in military protocols is too complex to adequately inform the average military reader (Flesch-Kincaid level of 6).

Method and Procedures

Mader and Playe (1997) explored the readability of consent forms used in emergency medicine research. Their method served as a foundation for this initial pilot study. Mader and Playe chose a descriptive, two-factor research design to compare the readability indices of consent forms (n = 94) over three categories separated by level of risk. The researchers analyzed the means of the groups with ANOVA and the Kruskal-Wallis test. They reported that the readability necessary to understand the consent forms rose as the risk of the study changed. Their results were significant ($p = .03$).

Procedures

For the pilot and full study, the Chiefs of the Clinical Investigation Regulatory Office of the Army and the Air Force were contacted and asked to provide copies of consent forms over

the time period 1999 – 2003. We anticipated some consent forms would be provided in hard copy while others would be in Adobe Acrobat™ (.pdf) files. Hard copy consent forms were scanned using commercial optical character recognition (OCR) software. Graphical and character mistakes from the scanning process were manually corrected to reflect the original. Using Microsoft Word™ (2003), each consent form was evaluated for readability and the results printed. For the full study, the consent forms and their readability scores were sorted into two categories of research based on risk: minimal risk (n = 30), and greater than minimal risk (n = 30). The pilot study analyzed only 10 consent forms from each risk category. Results were summarized into tabular format. Variables from the categories were compared with analysis of variance (ANOVA).

Consent form readability variables include the Flesch Reading Ease and Flesch-Kincaid grade level, words, characters, paragraphs, and sentences per document. All descriptive results were tabulated and sorted based on readability. Results were analyzed with SPSS version 12.0.

Expected Findings and Utility of Results

We expected to reject the alternate hypothesis and accept the null. From observation, military consent forms do not differ from those of the civilian sector. The results are quite predictable. What is more important is the implication of this conclusion.

In *Canterbury v. Spence* (1972), Judge Robinson briefly discussed the need for expert testimony in nondisclosure litigation. Despite the need for experts, it was lay testimony that "competently established a physician's failure to disclose particular risk information, the patient's lack of knowledge of this risk, and the adverse consequences following the treatment." Judge Robinson's statement could be addressed through the readability of consent forms. If a consent form is worded in a manner that a lay person can understand, then the participant is more

likely to understand the risks, benefits, and alternatives to the research procedures. Despite this landmark case in 1972, primary researchers have continued to author consent forms far beyond the ability of the average reader. Why has it not caused a widespread problem? If oncology consent forms, some in excess of 20 pages, are regularly signed, does the readability of these complex forms really matter? What is the efficacy of the consent form in deference to the desperation of the subject?

This document has a readability score of 12.0 (see Appendix C).

Data

Figure 1 illustrates a typical readability statistic report provided by Microsoft Word™. The Flesch Reading Ease provides an integer value commensurate to the ease of reading. The higher the number, the easier the document is to read. The Flesch Reading Ease calculates its result datum using a mathematical function as follows:

$[206.835 - (1.015 \times (avgwords / sentence)) - (84.6 \times (avgsyllables / word))]$.

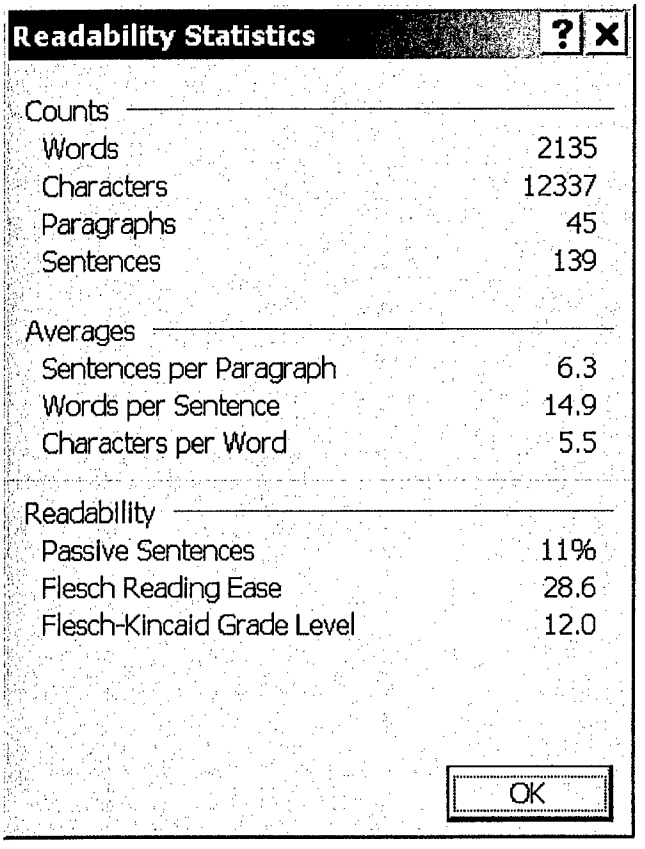
The reading ease score of 28.6 is suboptimal. The aim is to maximize the score with a score of 65 interpreted as "plain English" (Gnome, 2003). As the score approaches 100, the ease of reading improves.

The Flesch-Kincaid Grade Level calculates a similar score, but instead of a reading ease, it presents the school grade that an individual would need to have reached to understand the document. The Flesch-Kincaid Grade Level calculates this datum as follows:

$[.39 \times (avgwords / sentence) + (11.8 \times (avgsyllables / word) - 15.59]$. The Flesch-Kincaid Grade Level of 12.0 means that in order for a reader to understand this document, he/she would have to be, at a minimum, a high school graduate. Because the Flesch-Kincaid Grade Level ranges from

1 to 12.0, a document written beyond the high school graduate education level would still be represented by a 12.0.

Figure 1: Readability statistics for a document, calculated by Microsoft Word™.



Source: Microsoft Office Word, 2003 (11.6113.5703).

Types of Data

The Flesch-Kincaid readability criterion provides two calculations: Flesch Reading Ease and Flesch-Kincaid Grade Level. There are four types of data, and these are depicted in Figure 2. Because each readability criterion provides data in tenths, the characteristics of the data match the interval classification. As a result, parametric tests may be used.

Figure 2: Types of Data

Type of Data	Characteristics of Data	Basic Empirical Operation	Example
Nominal	Classification but no order, distance, or origin	Determination of equality	Gender (male, female)
Ordinal	Classification and order but no distance or unique origin	Determination of greater or lesser value	Doneness of meat (well, medium well, medium rare, rare)
Interval	Classification, order, and distance but no unique origin	Determination of equality of intervals or differences	Temperature in degrees
Ratio	Classification, order, distance, and unique origin	Determination of equality of ratios	Age in years

Source: Cooper and Schindler (2003, p. 233).

Probability Sampling Design

Cooper and Schindler (2003) list five designs for probability sampling. Appendix B illustrates these designs. Mader and Playe's (1997) design for this study included three groups. We modified this study design to delineate two risk categories in accordance with the Common Rule and Army Regulation 40-38. The probability sample fits the *stratified* description. We divided our protocols into groups, or strata, and maintained equal sample sizes in each group.

Statistical Techniques

With interval data and parametric tests, the tests available to evaluate the data are the *t* or *Z*-test (parametric) and analysis of variance (ANOVA). Figure 3 depicts a method by which researchers can select the appropriate test for statistical analysis. This study used this figure to determine an appropriate statistical technique to evaluate the data.

Figure 3: Statistical Techniques

Measurement Level	One-Sample Case	Two-Samples Case		k-Samples Case	
		Related Samples	Independent Samples	Related Samples	Independent Samples
Nominal	• Binomial • χ^2 One-sample	• McNemar	• Fisher exact test • χ^2 Two-samples test	• Cochran Q	• χ^2 for *k* samples
Ordinal	• Kolmogorov-Smirnov one-sample test • Runs test	• Sign test • Wilcoxon matched-pairs test	• Median test • Mann-Whitney U • Kolmogorov-Smirnov • Wald-Wolfowitz	• Friedman two-way ANOVA	• Median extension • Kruskal-Wallis one-way ANOVA
Interval and ratio	• *t*-test • *Z* test	• *t*-test for paired samples	• *t*-test • *Z* test	• Repeated-measures ANOVA	• One-way ANOVA • *n*-way ANOVA

Source: Cooper and Schindler (2003, p. 534).

General Linear Model Multivariate Analysis

Because our data are interval in nature, we can choose between nonparametric tests or stronger parametric tests, depending on the distribution of the data. If our data is normally distributed, the general linear model (GLM) multivariate analysis (version 12) is an appropriate parametric test. The following is a description of the GLM multivariate analysis, as explained in the help file of SPSS (version 12.0).

> The GLM multivariate procedure provides regression analysis and analysis of variance for multiple dependent variables by one or more factor variables or covariates. The factor variables divide the population into groups. Using this general linear model procedure, you can test null hypotheses about the effects of factor variables on the means of various groupings of a joint distribution of dependent variables. You can investigate interactions between factors as well

as the effects of individual factors. In addition, the effects of covariates and covariate interactions with factors can be included. For regression analysis, the independent (predictor) variables are specified as covariates.

Both balanced and unbalanced models can be tested. A design is balanced if each cell in the model contains the same number of cases. In a multivariate model, the sums of squares due to the effects in the model and error sums of squares are in matrix form rather than the scalar form found in univariate analysis. These matrices are called SSCP (sums-of-squares and cross-products) matrices. If more than one dependent variable is specified, the multivariate analysis of variance using Pillai's trace, Wilks' lambda, Hotelling's trace, and Roy's largest root criterion with approximate F statistic are provided as well as the univariate analysis of variance for each dependent variable. In addition to testing hypotheses, GLM Multivariate produces estimates of parameters.

We expect the means of the two groups to be normally distributed for all the dependent variables associated with each risk category.

Pilot Study

This pilot study was conducted jointly with CPT Kruse under the direction of COL Lee Briggs, Preceptor for the residency portion of the Army-Baylor Program. The results are mirrored identically in CPT Kruse's study. The purpose of conducting the pilot study was to verify the appropriateness of the procedure intended for use in both main studies, one of Army consent forms and the other of Air Force consent forms. At the time this pilot study was conducted, CPT Kruse and I were only granted access to the Army's consent forms. The lessons

learned from the pilot study were incorporated into the main study to further increase validity and reliability of the results.

Data (n = 20) for medical research studies were entered into SPSS coding groups as dichotomous variables (1 or 0), and recording integer output for the Flesch Reading Ease and the Flesch-Kincaid Grade Level. Table 1 displays the data.

Table 1. Sample Data Set

Consent Forms	Risk Level (GTMR = 0, MR = 1)	Words	Characters	Paragraphs	Sentences	Avg Sentences per Paragraphs	Avg Words per Sentences	Avg Characters per Word	Passive Sentences	Flesch Reading ease	Flesch-Kincaid Grade Level
1	0	3887	19891	116	157	4.3	20.5	4.8	0.35	41.9	12.0
2	0	2929	14959	75	138	3.3	19.1	4.9	0.28	39.4	12.0
3	0	4343	22061	274	182	2.1	18.4	4.8	0.29	46.7	11.1
4	0	3997	19762	97	169	3.5	20.2	4.7	0.45	46.6	11.6
5	0	6528	33364	307	292	2.1	18.6	4.8	0.33	48.5	10.8
6	0	5288	27314	183	215	2.5	21.2	4.9	0.35	42.4	12.0
7	0	1881	9908	51	78	2.8	21.4	5.1	0.35	33.8	12.0
8	0	7815	39329	272	255	2.7	24.4	4.8	0.37	41.5	12.0
9	0	1631	8431	48	58	2.0	25.6	5.0	0.41	33.6	12.0
10	0	1596	8647	54	67	3.5	20.6	5.1	0.23	32.4	12.0
11	1	1,604	8753	52	70	2.4	21.2	5.2	0.00	28.4	12.0
12	1	918	4987	53	35	1.8	21.4	5.0	0.37	30.6	12.0
13	1	918	5067	53	32	2.0	21.5	5.2	0.21	28.6	12.0
14	1	1210	6731	30	44	3.6	22.0	5.0	0.31	28.9	12.0
15	1	2783	14749	86	113	3.6	19.8	4.9	0.35	39.9	12.0
16	1	1997	10742	56	70	2.4	23.9	5.1	0.20	27.4	12.0
17	1	1743	9265	64	67	2.9	18.9	5.0	0.25	37.0	12.0
18	1	2369	12423	79	88	2.9	19.9	4.9	0.43	38.8	12.0
19	1	1755	9087	63	71	1.9	22.4	4.9	0.16	35.4	12.0
20	1	1115	5926	40	43	2.3	22.8	5.2	0.13	30.0	12.0

Results of Pilot Study

The results of the pilot study are summarized in Table 2. Ten minimal risk (minimal risk variable equal to one) and ten greater than minimal risk (minimal risk variable equal to zero) consent forms were analyzed. As depicted in Table 2, the mean number of words for minimal risk and greater than minimal risk consent forms were 1,641.20 ±623.03 and 3,989.50±2101.9, respectively. The mean number of words per sentence for the minimal and greater than minimal risk consent forms was 21.38±1.52 and 21.0±2.36, respectively. Overall, the mean reading ease score was 36.59±6.66 for all 20 forms. The reading scores for minimal risk consent forms was 32.5±4.76, and the mean reading score for greater than minimal risk consent forms was 40.68±5.82. The overall mean Flesch-Kincaid grade level was calculated at11.88±.33 for all 20 forms while the mean grade level for minimal risk was 12.0±0 and 11.7±.45 for greater than minimal risk. As depicted in Table 3, results showed seven items of significance. Each dependent variable increased along with risk. The following variables were significant at $P <$ .01: words ($F = 11.47$), characters ($F = 11.08$), sentences ($F = 13.86$), and Flesch Reading Ease ($F = 11.83$). The following dependent variables were significant at $P < .05$: paragraphs ($F =$ 7.50), average characters per word ($F = 6.47$), and passive voice ($F = 4.86$).

Lessons Learned

At the beginning of the study, the Army, Navy, and Air Force were contacted about participating. Initially, neither the Navy nor the Air Force provided any consent forms. The Director, Clinical Investigation and Responsible Conduct of Research for the U.S. Navy responded to our request for consent forms with extreme trepidation. The contact explained that the author of each study would have to be contacted and give permission to analyze the study's consent form. Further, the director intended to redact all information about the origin of the

study, the principle investigator, and any contact information. After agreeing to these terms, the director still failed to provide any data. The Division Chief for Biomedical Research and Compliance Division for the U.S. Air Force appeared cooperative to our initial requests, but consent forms were not provided. As a result, the pilot study was conducted using only Army consent forms. A few Air Force consent forms arrived after the Army study was complete. It was decided to use service-specific consent forms in distinct studies.

The Mader and Playe (1997) study which we originally planned to model chose a descriptive, two-factor research design to compare the readability indices of consent forms (n = 94) over three categories separated by level of risk. The Common Rule and Army Regulation 40-38 only delineate risk into two categories: minimal risk and greater than minimal risk. Creating a third category would necessitate utilizing the opinions of IRB members and research experts to assist in separating the available consent forms into three instead of two risk categories. To eliminate any human bias or error, we chose to study the consent forms based on risk specifically defined in The Common Rule and Army regulations.

As noted previously, Microsoft Word™ computes the Flesch-Kincaid Grade Level between a range of 1.0 to 12.0. Because the Flesch-Kincaid Grade Level in this program does not calculate grade levels above 12.0, a document written beyond that level would still be represented by a 12.0. Expectations are for the grade level of many of the protocols analyzed in this study to exceed the maximum score of 12.0 grade level. Since this study seeks to determine the magnitude of the number of informed consent forms that exceed the 6^{th} grade level, the limitation of the measurement tool is acceptable for this study's purpose. If another software program were utilized to calculate the average number of syllables per word for each consent form, the Flesch-Kincaid Grade Level could be calculated manually to validate our notion. The

lower risk consent documents were coded as one and the higher risk as zero. This is counterintuitive. Coding for the main study (n = 60) was reversed: coding minimal risk as zero and greater than minimal risk as one.

Table 2. Descriptive Statistics

Minimal Risk		Words	Characters	Paragraphs	Sentences	Avg Sentences per Paragraphs	Avg Words per Sentences	Avg Characters per Word	Passive Sentences	Flesch Reading ease	Flesch-Kincaid Grade Level
0	Mean	3989.50	20366.60	147.70	161.10	2.8800	21.0000	4.8900	.3410	40.6800	11.7500
	N	10	10	10	10	10	10	10	10	10	10
	Std. Deviation	2101.900	10533.893	102.716	79.090	.75542	2.35938	.13703	.06367	5.82348	.44535
1	Mean	1641.20	8773.00	57.60	63.30	2.5800	21.3800	5.0400	.2410	32.5000	12.0000
	N	10	10	10	10	10	10	10	10	10	10
	Std. Deviation	623.034	3224.843	16.608	25.404	.65286	1.51570	.12649	.12853	4.76049	.00000
Total	Mean	2815.35	14569.80	102.65	112.20	2.7300	21.1900	4.9650	.2910	36.5900	11.8750
	N	20	20	20	20	20	20	20	20	20	20
	Std. Deviation	1930.748	9636.336	85.233	76.064	.70420	1.93986	.14965	.11126	6.66388	.33226

Table 3. Results based on category of risk

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Risk_MR_1	Words	27572564.5	1	27572564.45	11.474	.003
	Characters	672057805	1	672057804.8	11.075	.004
	Paragraphs	40590.050	1	40590.050	7.498	.014
	Sentences	47824.200	1	47824.200	13.861	.002
	avg_spp	.450	1	.450	.903	.355
	avg_wps	.722	1	.722	.184	.673
	avg_cpw	.113	1	.113	6.470	.020
	passive	.050	1	.050	4.860	.041
	Reading_Ease	334.562	1	334.562	11.827	.003
	Grade_level	.313	1	.313	3.151	.093

Main Study of Army Consent Forms

The main study incorporated the lessons learned from the pilot study and data collected from 40 more consent forms, 20 additional minimal risk and 20 additional greater than minimal risk, for a total of 60 consent forms. Statistical analysis was conducted utilizing SPSS version 12.0.

Table 4. Full Army Data Set

Consent Forms	Risk (GTMR=1, MR=0)	Words	Characters	Paragraphs	Sentences	Avg Sentences per Paragraphs	Avg Words per Sentences	Avg Characters per Word	Passive Sentences	Flesch Reading ease	Flesch-Kincaid Grade Level	Syllables	New Flesch-Kincaid Grade Level
1	1	3887	19891	116	157	4.3	20.5	4.8	0.35	41.9	12.0	6433	11.93
2	1	2929	14959	75	138	3.3	19.1	4.9	0.28	39.4	12.0	4841	11.36
3	1	4343	22061	274	182	2.1	18.4	4.8	0.29	46.7	11.1	6920	10.39
4	1	3997	19762	97	169	3.5	20.2	4.7	0.45	46.6	11.6	6346	11.02
5	1	6528	33364	307	292	2.1	18.6	4.8	0.33	48.5	10.8	10758	11.11
6	1	5288	27314	183	215	2.5	21.2	4.9	0.35	42.4	12.0	8796	12.31
7	1	1881	9908	51	78	2.8	21.4	5.1	0.35	33.8	12.0	8243	44.47
8	1	7815	39329	272	255	2.7	24.4	4.8	0.37	41.5	12.0	12721	13.13
9	1	1631	8431	48	58	2.0	25.6	5.0	0.41	33.6	12.0	2327	11.23
10	1	1596	8647	54	67	3.5	20.6	5.1	0.23	32.4	12.0	2815	13.26
11	1	1024	5660	35	40	3.0	21.9	5.3	0.00	25.4	12.0	1911	14.97
12	1	1680	9292	77	69	1.9	21.0	5.1	0.28	31.3	12.0	2945	13.29
13	1	2562	12921	138	101	2.4	19.7	4.7	0.30	46.7	11.5	4089	10.93
14	1	4748	24330	153	204	2.9	21.1	4.9	0.36	39.8	12.0	7789	12.00
15	1	3986	20660	136	176	3.5	18.4	4.8	0.39	46.8	10.6	6496	10.82
16	1	4251	22276	136	153	1.9	23.9	4.9	0.30	42.2	12.0	7096	13.43
17	1	1392	7492	61	54	2.2	21.0	5.0	0.27	35.9	12.0	2491	13.72
18	1	4217	20852	111	183	3.5	19.9	4.6	0.37	49.9	10.6	6675	10.85
19	1	3738	19057	186	140	2.2	22.0	4.7	0.33	43.8	12.0	5999	11.93
20	1	3959	19386	124	180	2.6	19.0	4.7	0.39	51.3	10.7	6180	10.24
21	1	3616	18623	168	117	2.3	21.2	5.0	0.21	36.7	12.0	6056	12.44
22	1	5438	26814	175	244	2.0	20.2	4.7	0.49	50.1	11.1	8691	11.15
23	1	3703	18478	155	139	2.8	22.7	4.7	0.19	43.2	12.0	5814	11.79
24	1	4217	21056	181	145	2.7	18.7	4.7	0.26	46.8	11.2	6942	11.13
25	1	3941	20188	145	167	2.9	19.3	4.8	0.34	44.2	11.7	6489	11.37
26	1	2307	12117	48	98	3.5	22.4	5.1	0.23	33.9	12.0	3942	13.31
27	1	5365	27360	142	208	2.9	22.1	4.8	0.27	40.2	12.0	8816	12.42
28	1	1414	7608	68	63	2.1	19.8	5.1	0.17	32.6	12.0	2508	13.06
29	1	4490	23337	167	160	2.4	23.1	4.9	0.15	38.4	12.0	7551	13.26
30	1	4576	23160	178	169	2.6	21.6	4.7	0.24	46.3	11.9	7443	12.03

Table 4 (continued). Full Army Data Set

Consent Forms	Risk (GTMR=1, MR=0)	Words	Characters	Paragraphs	Sentences	Avg Sentences per Paragraphs	Avg Words per Sentences	Avg Characters per Word	Passive Sentences	Flesch Reading ease	Flesch-Kincaid Grade Level	Syllables	New Flesch-Kincaid Grade Level
1	0	1604	8753	52	70	2.4	21.2	5.2	0.00	28.4	12.0	2861	13.73
2	0	918	4987	53	35	1.8	21.4	5.0	0.37	30.6	12.0	1661	14.11
3	0	918	5067	53	32	2.0	21.5	5.2	0.21	28.6	12.0	1670	14.26
4	0	1210	6731	30	44	3.6	22.0	5.0	0.31	28.9	12.0	2100	13.47
5	0	2783	14749	86	113	3.6	19.8	4.9	0.35	39.9	12.0	4720	12.14
6	0	1997	10742	56	70	2.4	23.9	5.1	0.20	27.4	12.0	3515	14.50
7	0	1743	9265	64	67	2.9	18.9	5.0	0.25	37.0	12.0	3005	12.12
8	0	2369	12423	79	88	2.9	19.9	4.9	0.43	38.8	12.0	3995	12.07
9	0	1755	9087	63	71	1.9	22.4	4.9	0.16	35.4	12.0	2977	13.16
10	0	1115	5926	40	43	2.3	22.8	5.2	0.13	30.0	12.0	1954	13.98
11	0	2374	12461	88	97	3.3	18.5	4.9	0.50	40.4	11.9	4033	11.67
12	0	1430	7696	53	64	2.2	19.9	5.2	0.17	29.4	12.0	2539	13.12
13	0	1950	9899	82	75	3.1	18.4	4.8	0.26	46.3	10.9	3205	10.98
14	0	2015	11069	74	73	2.4	21.1	5.0	0.34	33.5	12.0	3569	13.54
15	0	1468	8060	35	60	3.0	22.2	5.3	0.40	28.5	12.0	2606	14.02
16	0	1318	7200	47	47	1.8	25.3	5.2	0.17	18.9	12.0	2359	15.40
17	0	2206	11463	86	85	2.9	18.9	4.9	0.36	39.5	12.0	3729	11.73
18	0	2114	11874	72	64	2.2	26.8	5.2	0.25	25.4	12.0	3818	16.17
19	0	3385	17715	99	131	3.4	21.5	4.9	0.31	38.7	12.0	5370	11.51
20	0	1673	8768	74	63	2.4	17.9	4.9	0.41	36.5	12.0	2850	11.49
21	0	2627	14126	82	98	3.1	20.9	5.0	0.26	36.0	12.0	4464	12.61
22	0	2391	12673	86	90	2.5	20.5	5.0	0.24	36.0	12.0	4060	12.44
23	0	2154	11827	79	70	2.5	23.0	5.0	0.27	33.6	12.0	3815	14.28
24	0	2098	11179	76	85	3.1	19.2	4.9	0.32	39.4	12.0	3552	11.88
25	0	1779	9465	54	67	2.4	23.7	5.1	0.23	29.0	12.0	3106	14.25
26	0	1738	9217	63	65	2.2	22.2	5.0	0.21	30.0	12.0	3116	14.22
27	0	1760	9173	48	79	2.9	20.4	4.9	0.27	35.4	12.0	3029	12.67
28	0	6091	31178	182	264	3.8	18.5	4.8	0.26	46.3	11.1	9872	10.75
29	0	4398	22283	285	182	2.1	18.4	4.8	0.17	46.8	11.0	6985	10.33
30	0	2056	10610	75	77	2.7	20.3	4.8	0.28	43.9	11.7	3429	12.01

Results of Main Study

The results of the main study are summarized in Tables 5 and 6. The mean number of words for minimal risk and greater than minimal risk consent forms were 2,114.57 ± 1032.70 and 3,683.97 ± 1,607.72 respectively. Similarly, for the dependent variables of characters, paragraphs, sentences, average sentences per paragraph, and passiveness, the mean values were larger for greater than minimal risk than for minimal risk consent forms. However, the mean values for the dependent variables of average words per sentence, average characters per word, Flesch Reading Ease, and Flesch-Kincaid Grade Level were all larger for minimal risk consent forms than greater than minimal risk consent forms. Overall, the mean Flesch Reading Ease and the Flesch-Kincaid Grade Level score for all 60 forms was 37.65 ± 7.34 and 11.79±.42 respectively. As illustrated in Table 6, five of the ten dependent variables were significant ($P<.01$) based on level of risk which included number of words ($F = 20.24$), number of characters ($F = 19.14$), number of paragraphs ($F=14.322$), number of sentences ($F=20.85$), and Flesch Reading Ease ($F=14.23$) demonstrating a direct relationship with level of risk.

Table 5. Descriptive Statistics for Main Study

risk_cat		words	char	parag	sentence	avg_s_pr	avg_w_s	avg_c_w	passive	fl_re	fk_gl
MR	Mean	2114.57	11188.87	77.20	82.30	2.6600	21.0467	5.0000	.2697	34.6167	11.8867
	Std. Deviation	1032.700	5210.382	47.781	45.079	.55436	2.15130	.14384	.10179	6.72858	.30708
	N	30	30	30	30	30	30	30	30	30	30
GTMR	Mean	3683.97	18811.10	135.37	147.37	2.7033	20.9667	4.8700	.2983	41.0767	11.6933
	Std. Deviation	1607.719	7994.846	69.313	63.720	.60429	1.81076	.17050	.09798	6.53414	.49126
	N	30	30	30	30	30	30	30	30	30	30
Total	Mean	2899.27	14999.98	106.28	114.83	2.6817	21.0067	4.9350	.2840	37.8467	11.7900
	Std. Deviation	1555.912	7715.690	65.907	63.804	.57534	1.97182	.16957	.10011	7.33816	.41770
	N	60	60	60	60	60	60	60	60	60	60

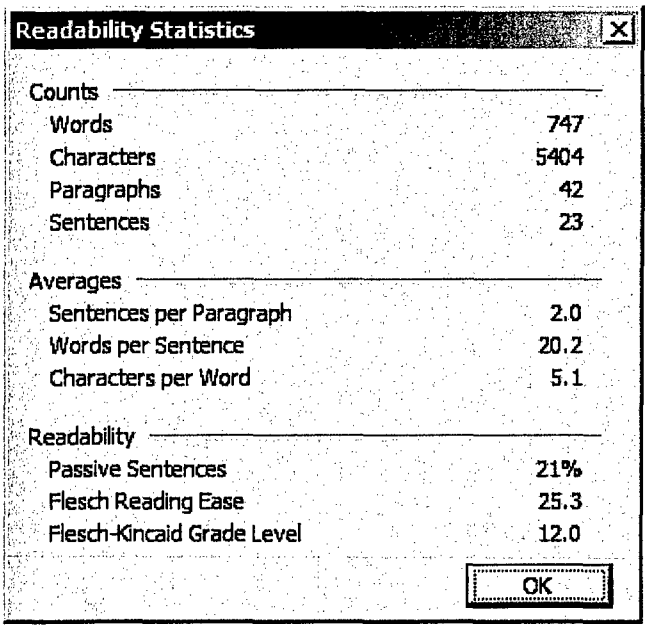
Table 6. Analysis of Variance for Main Study

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
risk_cat	words	36945245.4	1	36945245.40	20.237	.000
	char	871476615	1	871476614.8	19.140	.000
	parag	50750.417	1	50750.417	14.322	.000
	sentence	63505.067	1	63505.067	20.848	.000
	avg_s_pr	.028	1	.028	.084	.773
	avg_w_s	.096	1	.096	.024	.877
	avg_c_w	.254	1	.254	10.189	.002
	passive	.012	1	.012	1.235	.271
	fl_re	625.974	1	625.974	14.232	.000
	fk_gl	.561	1	.561	3.341	.073

Discussion

The primary difference between the Mader and Playe (1997) study and this study on readability of Army informed consent documents is the number of risk groups studied. Mader and Playe used three risk categories, while this study used only two categories as defined by The Common Rule and AR 40-38; minimal risk and greater than minimal risk. While Mader and Playe found significance between reading ease and reading level based on three risk categories, in this study, significance exists only for reading ease based on risk category. The differing results may be a consequence of the consent form template required by the Army. The human research subject's agreement to participate in a research protocol must be documented using Department of the Army (DA) Form 5303-R. A copy of the template is provided in Appendix D. The readability statistics of the form alone yielded the results found in Figure 4. Because the template itself is already at a reading level of 12, authoring a consent form with extremely simplistic verbiage may not lower the reading level of the document to an acceptable level. The template may need to be reviewed for possible revision if making human research consent forms easier to read is a goal for the Army. A subsequent study may seek to separate the specific verbiage of the consent form from that of the template to determine if the protocol-specific language is at an acceptable reading level.

Figure 4. Readability scores for DA Form 5303-R



This study found significance (F=14.23, P<.01) between risk categories on the Flesch Reading Ease score between risk categories. Oddly, the greater than minimal risk consent forms were found to be easier to read than the minimal risk consent forms. The mean reading ease scores were 41.08±6.53 and 34.62±6.73 respectively. This is the result of the researchers using simpler verbiage containing shorter sentences with few multi-syllabic words on the longer greater than minimal risk consent forms. Kruse (2005), performed the identical study presented here with United States Air Force consent forms (n=21). He found significance (p<.01) on four of the 10 dependent variables: number of words (F=13.51), number of characters (F=14.22), number of paragraphs (F=7.93), and number of sentences (F=15.00).

Because the two studies of Army and Air Force consent forms were performed similarly, the results can be compared. Army consent forms used passive voice 5% more frequently than Air Force consent forms (F=4.41, p<.05). When data from the two services were combined, significance was found (p<.01) between risk categories in four of the 10 dependent variables:

number of words (F=17.01), number of characters (F=16.96), number of paragraphs (F=9.10), and number of sentences (F=18.03). Finally, when comparing data between risk categories and services, only the Flesch-Kincaid Grade Level variable was significant (F=4.67, p<.05).

Regardless of the service component, Army or Air Force, the data reveal that human subject research consent forms are written at a level that exceeds that at which many volunteers can understand. The consent form is the building block upon which the foundation of informed consent is built. Without an understandable consent form, we can not be assured that human subject research participants will comprehend the risks, benefits, or alternatives of the research. Additionally, complex consent forms that mask risk, benefits, and alternatives in complex verbiage may inhibit participants from asking questions to clarify what they do not understand.

Limitations

While the results of this study confirm the findings of similar research in the civilian sector, (i.e., that consent forms are too difficult for the average human subject research participant to understand), there are several limitations of this study that could be addressed in future studies. First, only the Flesch-Kincaid Grade Level and Flesch Reading Ease readability tests were used to evaluate the readability of the ICDs. Use of several readability tests could strengthen the validity and reliability of the results. Second, as mentioned previously, variables other than average number of words per sentence and average syllables per word, as used in the Flesch-Kincaid Grade Level and Flesch Reading Ease calculations may, in fact, influence the readability of a document. For instance, the type font, size, and color, use of diagrams or pictures, spacing between paragraphs, or other document characteristics may affect the readability of a document. Third, because the ICD template mandated for use in Army research is at a reading grade level of 12 before any of the protocol-specific information (i.e., research

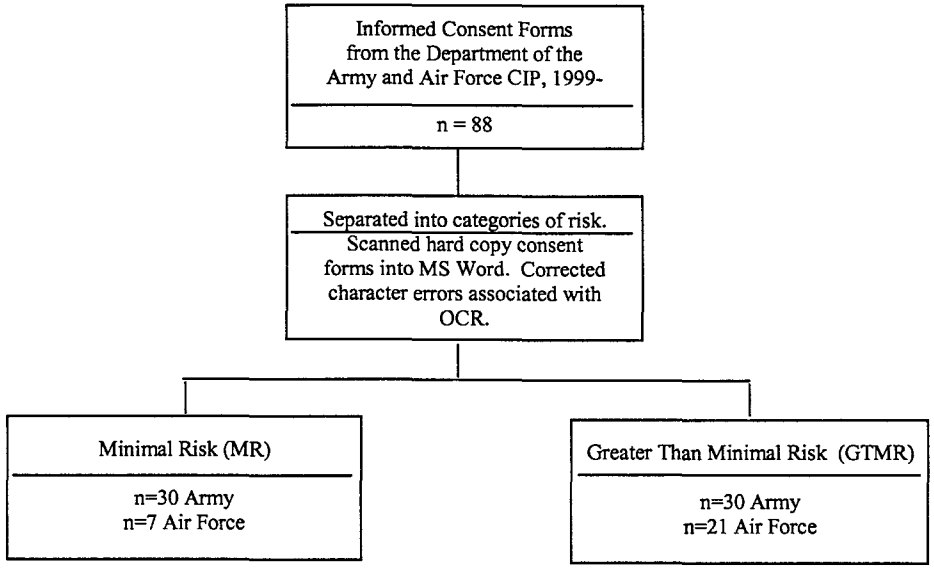
procedure, risks, benefits or alternatives) is added, the ICD author begins the process of writing an easy to read consent form at an extreme disadvantage. Lastly, medical literature, including ICDs, is fraught with long, multi-syllabic terminology of Latin origin. Normally these words are explained in the text of the consent form, but the use of the words themselves may adversely affect the readability score.

Conclusion

The results of this study are not surprising but frustrating. Despite the consistent results found in informed consent readability research that consent forms are far too difficult for the average prospective human research subject to understand, there has been no progress toward making them any easier to read. The major challenge for medical researchers within the Army now is to determine how to simplify consent forms to ensure that volunteers in human subject research can truly understand the risks, benefits, and alternatives of the research they may participate in. An excellent starting place may be to revise the informed consent template, DA Form 5303-R, required for use in human research protocols.

Appendix A

Research work flow diagram



Appendix B

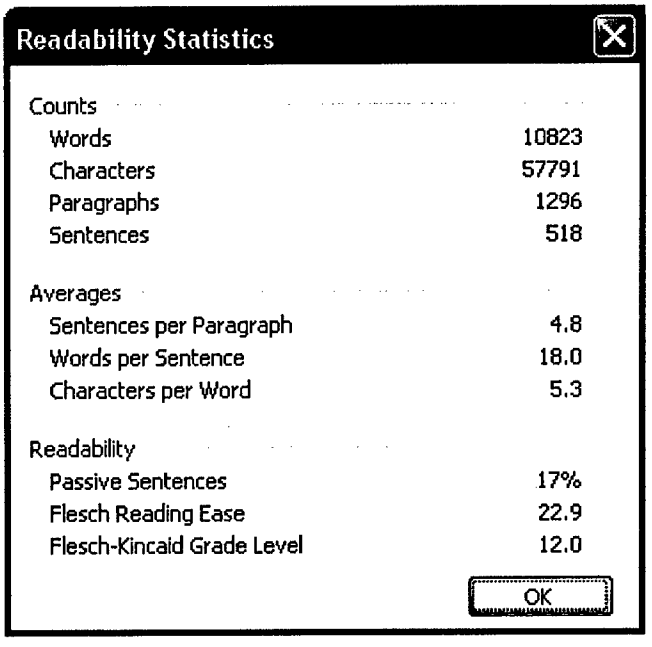
Comparison of Probability Sampling Designs

Type	Description	Advantages	Disadvantages
Simple random	Each population element has an equal chance of being selected into the sample. Sample drawn using random number table/generator.	Easy to implement with automatic dialing (random digit dialing) and with computerized voice response systems.	Requires a listing of population elements. Takes more time to implement. Uses larger sample sizes. Produces larger errors. Expensive.
Systematic	Selects an element of the population at a beginning with a random start and following the sampling fraction selects every *k*th element.	Simple to design. Easier to use than the simple random. Easy to determine sampling distribution of mean or proportion. Less expensive than simple random.	Periodicity within the population may skew the sample and results. If the population list has a monotonic trend, a biased estimate will result based on the start point.
Stratified	Divides population into subpopulations or strata and uses simple random on each strata. Results may be weighted and combined.	Researcher controls sample size in strata. Increased statistical efficiency. Provides data to represent and analyze subgroups. Enables use of different methods in strata.	Increased error will result if subgroups are selected at different rates. Expensive. Especially expensive if strata on the population have to be created.
Cluster	Population is divided into internally heterogeneous subgroups. Some are randomly selected for further study.	Provides an unbiased estimate of population parameters if properly done. Economically more efficient than simple random. Lowest cost per sample, especially with geographic clusters. Easy to do without a population list.	Often lower statistical efficiency (more error) due to subgroups being homogeneous rather than heterogeneous.
Double (sequential or multiphase)	Process includes collecting data from a sample using a previously defined technique. Based on the information found, a subsample is selected for further study.	May reduce costs if first stage results in enough data to stratify or cluster the population.	Increased costs if indiscriminately used.

Source: Cooper and Schindler (2003, p. 199).

Appendix C

Readability score for this study



Appendix D

Department of the Army Form 5303-R

VOLUNTEER AGREEMENT AFFIDAVIT

For use of this form, see AR 70-25 or AR 40-38; the proponent agency is OTSG.

PRIVACY ACT OF 1974

Authority:	10 USC 3013, 44 USC 3101, and 10 USC 1071-1087.
Principle Purpose:	To document voluntary participation in the Clinical Investigation and Research Program. SSN and home address will be used for identification and locating purposes.
Routine Uses:	The SSN and home address will be used for identification and locating purposes. Information derived from the study will be used to document the study; implementation of medical programs; adjudication of claims; and for the mandatory reporting of medical conditions as required by law. Information may be furnished to Federal, State and local agencies.
Disclosure:	The furnishing of your SSN and home address is mandatory and necessary to provide identification and to contact you if future information indicates that your health may be adversely affected. Failure to provide the information may preclude your voluntary participation in this investigational study.

PART A(1) - VOLUNTEER AFFIDAVIT

Volunteer Subjects in Approved Department of the Army Research Studies

Volunteers under the provisions of AR 40-38 and AR 70-25 are authorized all necessary medical care for injury or disease which is the proximate result of their participation in such studies.

I,_____, SSN_____,

having full capacity to consent and having attained _____ birthday, do hereby volunteer/give consent as legal

representative _____ to participate _____

_____

(Research study)

_____

under the direction _____

conducted at_____

(Name of Institution)

The implications of my voluntary participation/consent as legal representative; duration and purpose of the research study; the methods and means by which it is to be conducted; and the inconveniences and hazards that may reasonably be expected have been explained to me by

_____

I have been given an opportunity to ask questions concerning this investigational study. Any such questions were answered to my full and complete satisfaction. Should any further questions arise concerning my rights/the rights of the person I represent on study-related injury, I may contact

_____

at _____

(Name, Address and Phone Number of Hospital (Include Area Code))

I understand that I may at any time during the course of this study revoke my consent and withdraw/have the person I represent withdrawn from the study without further penalty or loss of benefits; however, I/the person I represent may be required (military volunteer) or requested *(civilian volunteer)* to undergo certain examination if, in the opinion of the attending physician, such examinations are necessary for my/the person I represent's health and well-being. My/the person I represent's refusal to participate will involve no penalty or loss of benefits to which I am/the person I represent is otherwise entitled.

_____

PART A(2) - ASSENT VOLUNTEER AFFIDAVIT *(MINOR CHILD)*

I,_____, SSN_____, having full

capacity to assent and having attained _____ birthday, do hereby volunteer _____

_____ to participate _____

_____

(Research Study)

_____

under the direction of_____

conducted at_____

(Name of Institution)

(Continue on Page 2)

DA FORM 5303-R, MAY 1989 PREVIOUS EDITIONS ARE OBSOLETE. USAPA V1 00

PART A(2) - ASSENT VOLUNTEER AFFIDAVIT *(MINOR CHILD) (Cont'd.)*

The implications of my voluntary participation; the nature, duration and purpose of the research study; the methods and means by which it is to be conducted; and the inconveniences and hazards that may reasonably be expected have been explained to me by

_____

I have been given an opportunity to ask questions concerning this investigational study. Any such questions were answered to my full and complete satisfaction. Should any further questions arise concerning my rights I may contact

_____

at _____

(Name, Address and Phone Number of Hospital (Include Area Code))

I understand that I may at any time during the course of this study revoke my consent and withdraw from the study without further penalty or loss of benefits; however, I may be requested to undergo certain examination if, in the opinion of the attending physician, such examinations are necessary for my health and well-being. My refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled.

_____

PART B - TO BE COMPLETED BY INVESTIGATOR

INSTRUCTIONS FOR ELEMENTS OF INFORMED CONSENT: *(Provide a detailed explanation in accordance with Appendix C, AR 40-38 or AR 70-25.)*

I do ☐ do not ☐ *(check one & initial)* consent to the inclusion of this form in my outpatient medical treatment record.

SIGNATURE OF VOLUNTEER	DATE	SIGNATURE OF LEGAL GUARDIAN *(If volunteer is a minor)*
PERMANENT ADDRESS OF VOLUNTEER	TYPED NAME OF WITNESS	
	SIGNATURE OF WITNESS	DATE

References

Albarelli, H. & Kelly, J. (2001, July 6). New evidence in Army's scientist's death [Electronic Version]. *World Net Daily.* Retrieved November 30, 2003 from http://www.worldnetdaily.com/news/aricle.asp?ARTICLE_ID=23489.

Arthur, V. (1995). Written patient information: a review of the literature. *Journal of advanced nursing.* 21(6), pp. 1081 – 1086.

Beauchamp, T. & Childress, J. (2001). *Principles of Biomedical Ethics* (5th ed.). New York: Oxford University Press.

Canterbury v. Spence, 464 F.2d 772 (D.C. Cir. 1972).

Cooper D., & Schindler P. (2003). Business research methods. New York: McGraw-Hill.

Ferguson, P.R., (2002). Patients' perceptions of information provided in clinical trials. *Research ethics.* 28, pp.45 – 48.

Glazer-Waldman, W., Hall, K., & Weiner, M., (1985). Patient education in a public hospital. *Nursing research.* 34(3), pp. 184 – 185.

Gnome (2003). *User and readability consideration for technical documentation.* Retrieved November 1, 2003, from http://developer.gnome.org/documents/style-guide/usability-readability.html.

Howard, J., & DeMets, D. How informed is informed consent? The BHAT experience. *Controlled clinical trials.* 1981(2), 287 – 303.

Johnson, K. (2004). Readability. Retrieved August 6, 2004 from http://www.timetabler.com/reading.html

Jones, J. (1993) *Bad blood*: *The Tuskegee syphilis experiment.* New York, The Free Press

Lee, P. (1999, January). Why literacy matters: Links between reading ability and health. Why literacy matters: Links between reading ability and health. *American Medical Association*, 117, 100-103.

Limitation On Use of Humans as Experimental Subjects, 10 U.S.C. 980. Retrieved April 4, 2005 from http://www.dtic.mil/biosys/downloads/title10.pdf

Luce, J. (2003). Is the concept of informed consent applicable to clinical research involving critically ill patients? *Critical Care Medicine.* 31(3) supplement, pp. S153 – S160.

Mader, T., & Playe, S. (1997). Emergency medicine research consent form readability assessment. *Annals of emergency medicine.* 29(4), pp. 534 – 539.

Matot, E., Pizov, R., & Sprung, C. (1998). Evaluation of institutional review board review and informed consent in publications of human research in critical care medicine. *Critical Care Medicine.* 26(9), pp. 1488 – 1489.

Microsoft. (2003). Microsoft Word 2000 (9.096926 SP-3). Belleview, WA: Microsoft Corporation.

National Defense, 32 C.F.R. § 219 (2004).

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979). *The Belmont report:.ethical principles and guidelines for the protection of human subjects of research.* Retrieved November 30, 2003 from http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm

Ott, B., & Hardie, T. (1997). Readability of advance directive documents. *The Journal of Nursing Scholarship.* 29(1), pp. 53 – 57.

Paasche-Orlow, M., Taylor, H., & Brancati, F. (2003). Readability standards for informed-consent forms as compared with actual readability. *The New England Journal of Medicine.* 348(8), pp. 721-726.

Public Welfare, 45 C.F.R. § 46 (2004).

Schloendorff v. Society of New York Hospital, 211 NY 125, 105 NE 92:1914.

Standler, R. (1997).*Nonconsensual Medical Experiments on Human Beings.* Retrieved November 30, 2003 from http://www.rbs2.com/humres.htm.

The United Nations General Assembly. (1948). *Universal Declaration of Human Rights.* Retrieved November 30, 2003 from http://www.un.org/Overview/rights.html.

U. S. Army: Clinical Investigation Program: Army Regulation 40-38. Washington, DC, Department of the Army, 1989.

U. S. Department of Health and Human Services (2004) *Institutional review board guidebook.* Retrieved November 8, 2003, from http://www.hhs.gov/ohrp/irb/irb_guidebook.htm .

United States v. Karl Brandt, et al., The Medical Case Trial, Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10. Nuremburg (1950).

University of Utah (2004a). *The Jewish chronic disease hospital study.* Retrieved November 30, 2003 from http://www.bioen.utah.edu/faculty/sri/LAB1_handout.htm.

University of Utah. (2004b). *The Smog readability formula.* Retrieved November 1, 2003, from http://www.med.utah.edu/pated/authors/readability.html

University of Utah. (2004c). *The Willowbrook study.* Retrieved November 1, 2003, from http://www.bioen.utah.edu/faculty/sri/LAB1_handout.htm.

Walmsley, S., & Allington, R.(1982). Reading abilities of elderly persons in relation to the difficulty of essential documents. *Gerontologist.* 22(1), pp. 36 – 38.

Woodward, B. (1999). Challenges to human subject protections in US medical research. *Journal of the American Medical Association*, 282, pp. 1947-1952.

Zucker, K., & Boyle, M. (2000). Health Law for Federal Sector Administrators (2 vol, 8th ed.). AMEDDC&S.